Exploring an evolution-centric statistical forecast technique for Western U.S. precipitation

Agniv Sengupta
JPL Postdoctoral Scholar
Advisor: Duane Waliser

Winter Outlook Workshop
November 17, 2021

Jet Propulsion Laboratory
California Institute of Technology

Outline

Part I: Motivation and hypothesis

Part II: Evolution-centric statistical forecast technique

Part III: Hindcast skill of winter precipitation over the western U.S.

Part IV: Experimental seasonal forecast for winter 2021-22

Part V: Future plans: Statistical and Machine Learning applications

Outline

Part I: Motivation and hypothesis

Part II: Evolution-centric statistical forecast technique

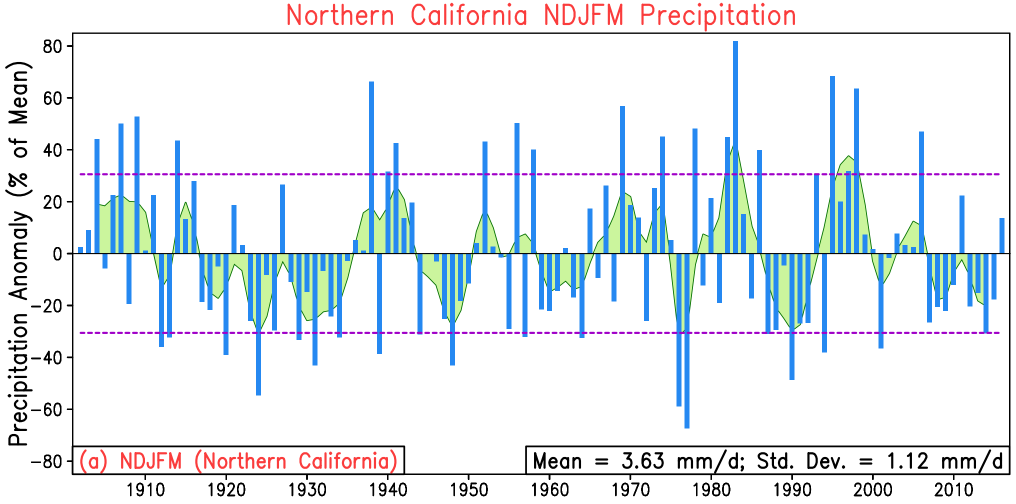
Part III: Hindcast skill of winter precipitation over the western U.S.

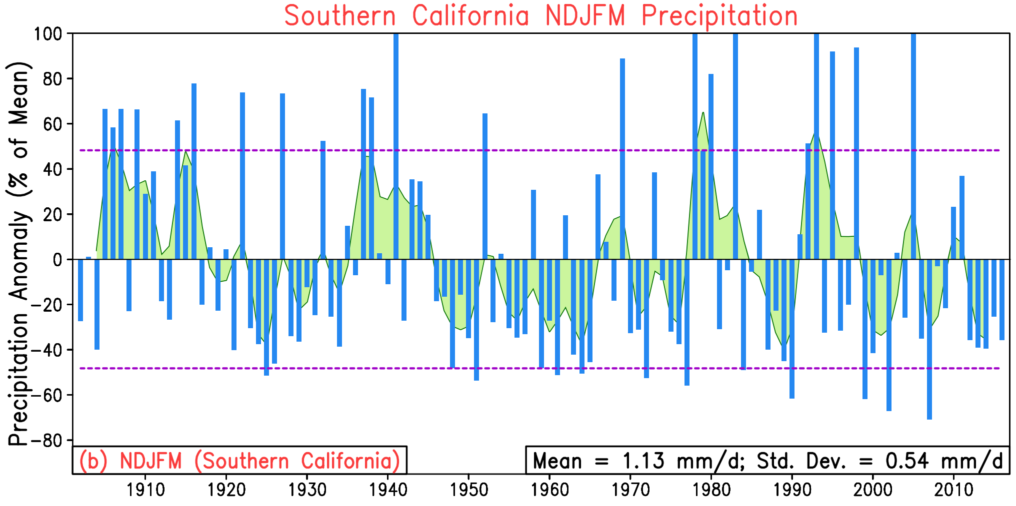
Part IV: Experimental seasonal forecast for winter 2021-22

Part V: Future plans: Statistical and Machine Learning applications

Hypothesis: Low-frequency variability in the forcing?

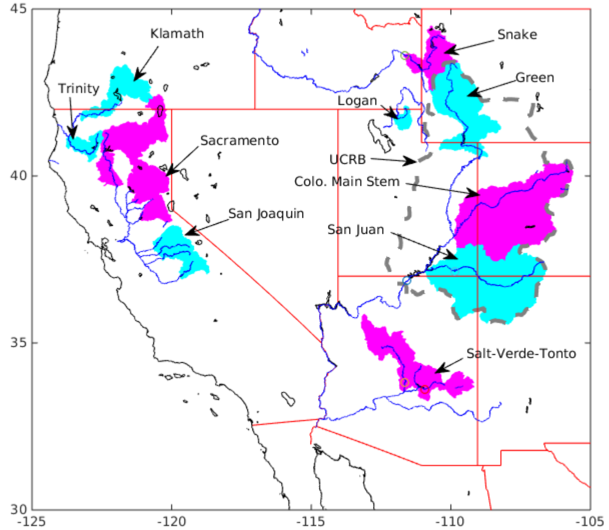
Observed Precipitation

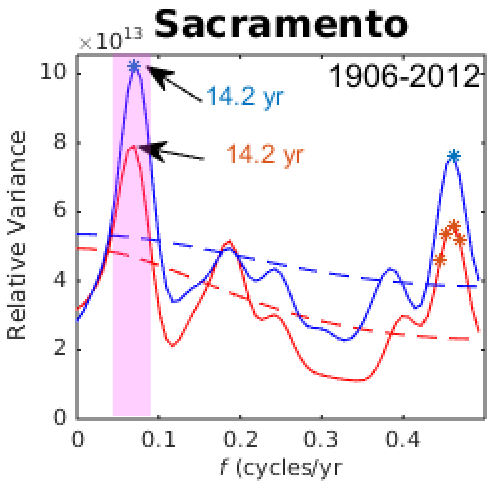


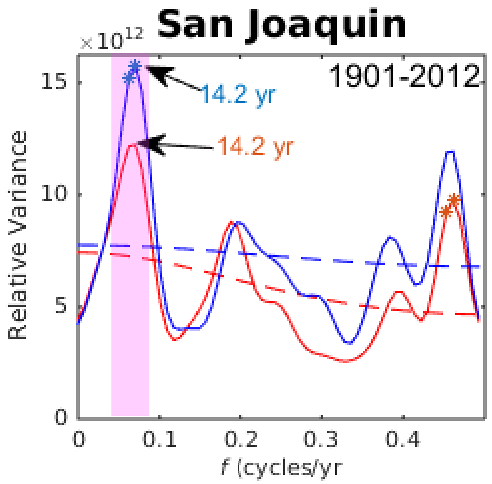


- Besides *year-to-year* fluctuations, observed precipitation reveals prominent role of *slowly changing* recurrences.

Tree-ring Records & Reconstructed Streamflow

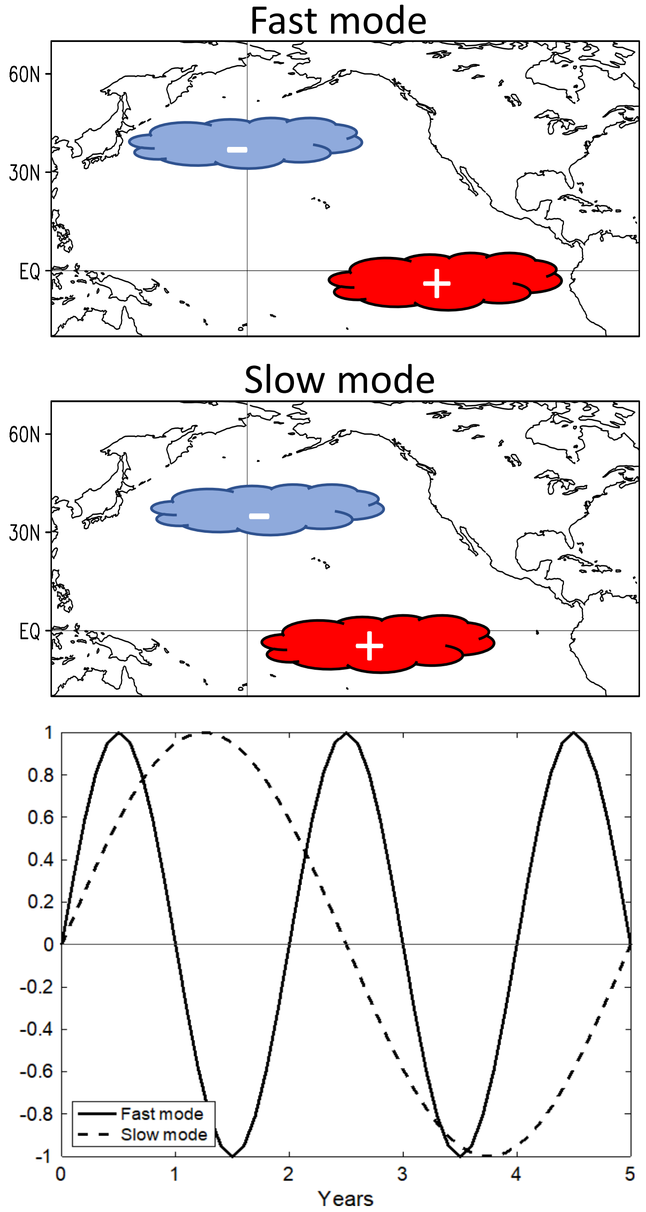




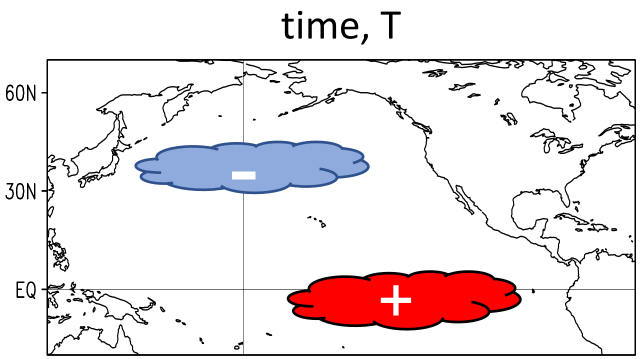


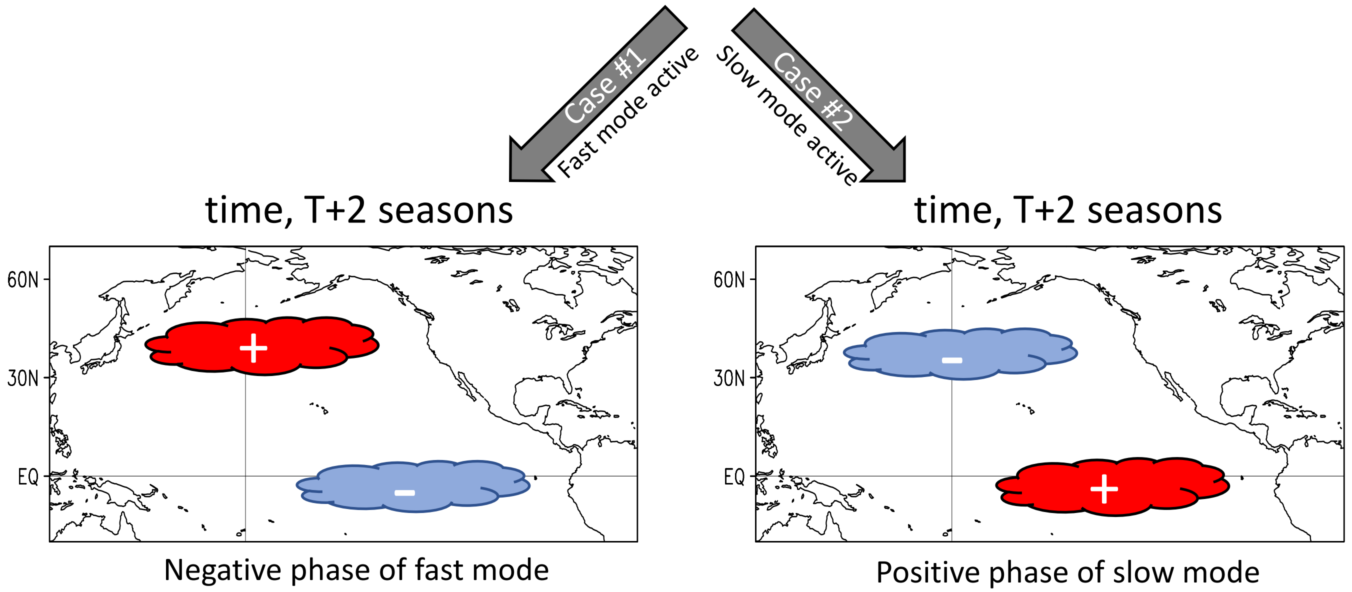
Reference: Meko, D.M., C.A. Woodhouse, and, E.R. Bigio. 2018. “Southern California Tree Ring Study.” Final Report to California DWR.

Sea surface temperature evolution: A key consideration in seasonal prediction



- *Which of the two modes is active — the 5-year mode, or, the 2-year one?*
- *Can we detect by simply looking at any single time point?*





Hence, it becomes critical to focus on the past ***multi-season*** structure of the predictor instead of just the present season for an accurate attribution of the dominant mode.

Outline

Part I: Motivation and hypothesis

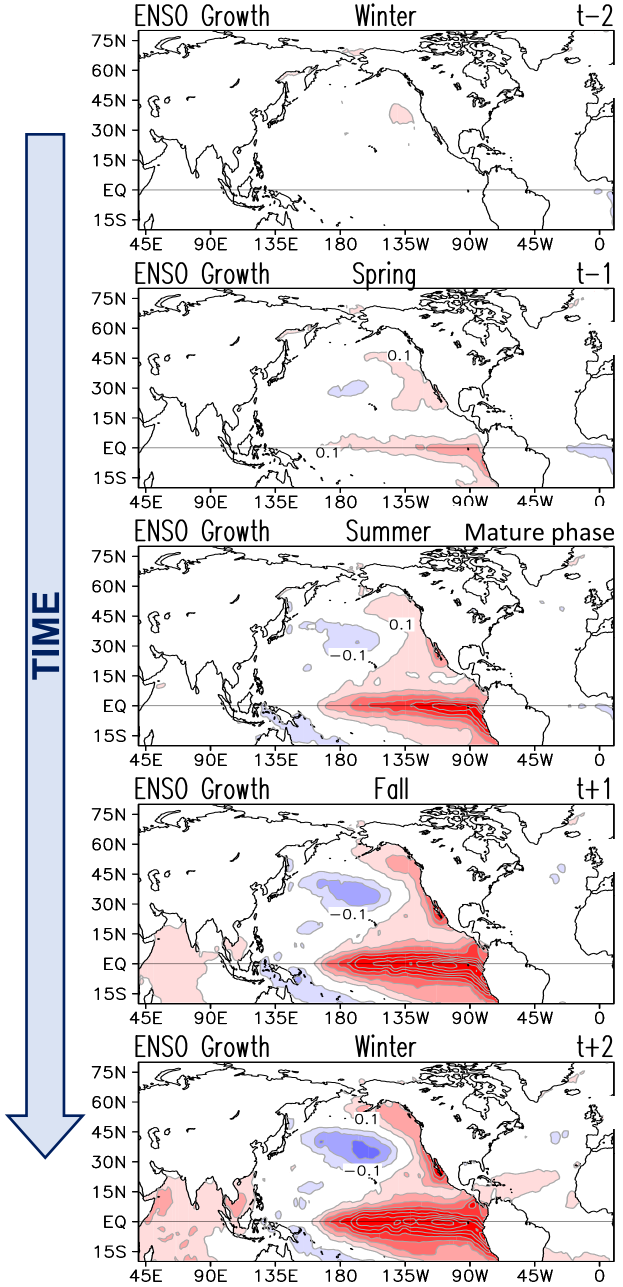
Part II: Evolution-centric statistical forecast technique

Part III: Hindcast skill of winter precipitation over the western U.S.

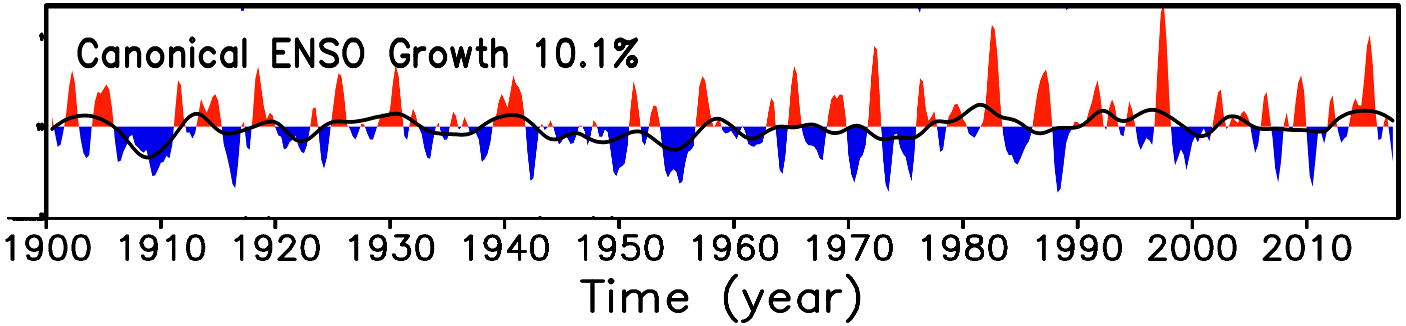
Part IV: Experimental seasonal forecast for winter 2021-22

Part V: Future plans: Statistical and Machine Learning applications

Predictors: *Multi-season sea surface temperatures*



- SSTs influence both regional and remote hydroclimate:
  - *Interannual variation*: ENSO impacts the North American hydroclimate, Indian summer monsoon
  - *Decadal variations*: Multi-year droughts, e.g., the 1930s ‘Dust Bowl’ over the Great Plains
- We analyze of *118 years* of observed, seasonal SST anomalies
- Technique: Extended-Empirical Orthogonal Function (extended-EOF) analysis
- Eleven modes of global SST variability (natural variability and secular trend) extracted
- Each comprises of a sequence of maps (or, the extended-EOF pattern), and its related time series (principal component). For example,



Canonical ENSO Growth extended-EOF mode

C.I. for SST anomalies = 0.1 K

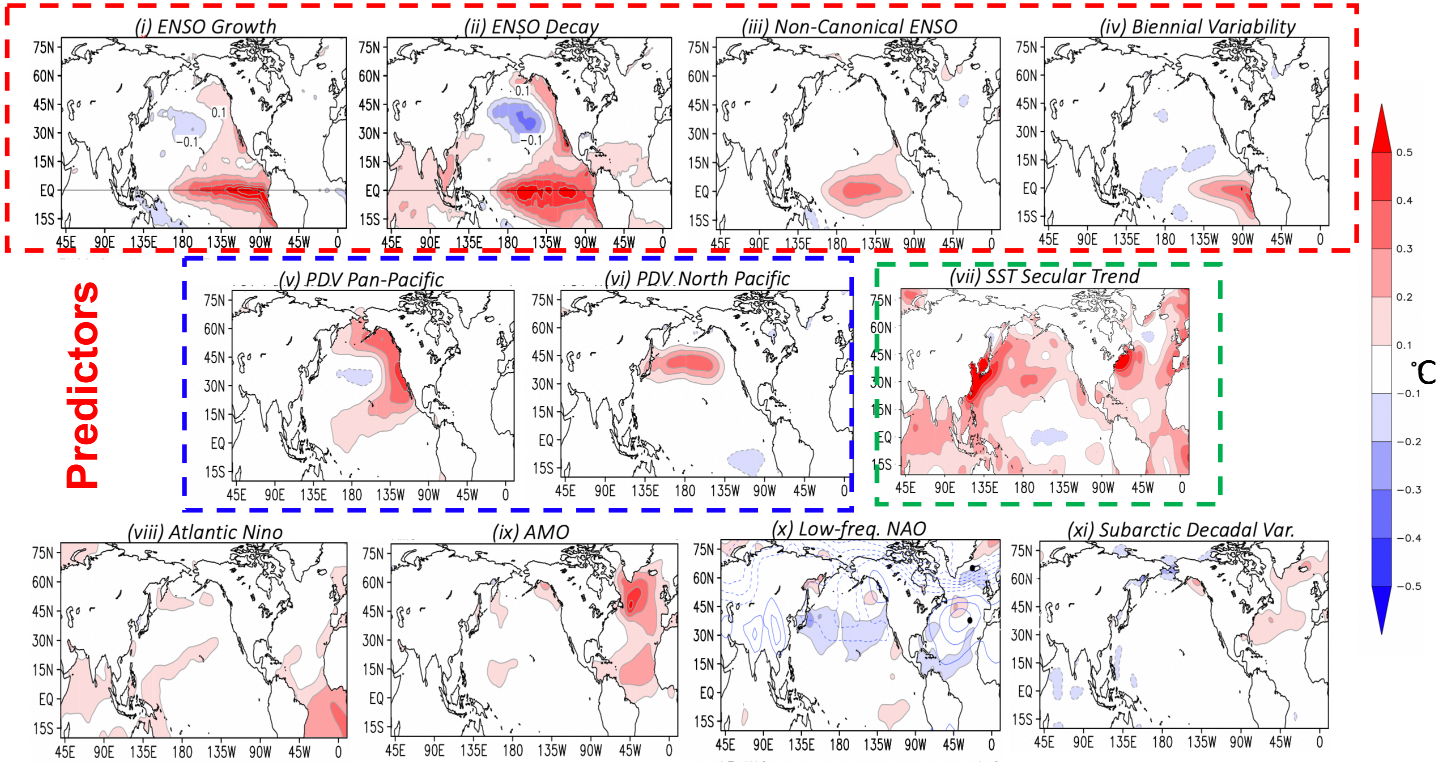
References:
- Nigam, S., A. Sengupta, and A. Ruiz-Barradas, 2020, *J. Climate*, 33(13), 5479-5505.
- Nigam, S. and A. Sengupta, 2021, *Geophysical Research Letters*, 48(3), https://doi.org/10.1029/2020GL091447.

Statistical forecast technique

This analysis leverages observational variables with large thermal inertia (e.g., SSTs) for skillful seasonal prediction.

Unique characteristics of our approach:

- use of multi-season, antecedent predictor information instead of utilizing just the preceding one season
- improved characterization of the evolution of the recurrent variations, i.e., both the *spatial and temporal* recurrence
- additional consideration of *lower-frequency* sources of natural variability in addition to interannual variability



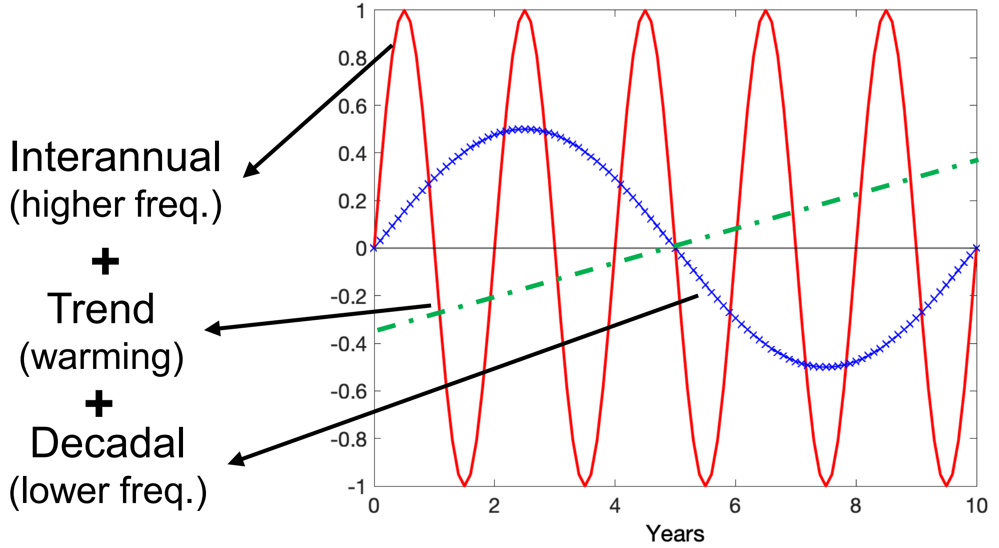


Figure *(left) Leading modes of global SST variability – ENSO (top row), Pacific Decadal Variability and Secular Trend (middle row), Atlantic modes (bottom row) informing seasonal prediction of precipitation. (right) Idealized representation of the frequency of SST modes*

Outline

Part I: Motivation and hypothesis

Part II: Evolution-centric statistical forecast technique

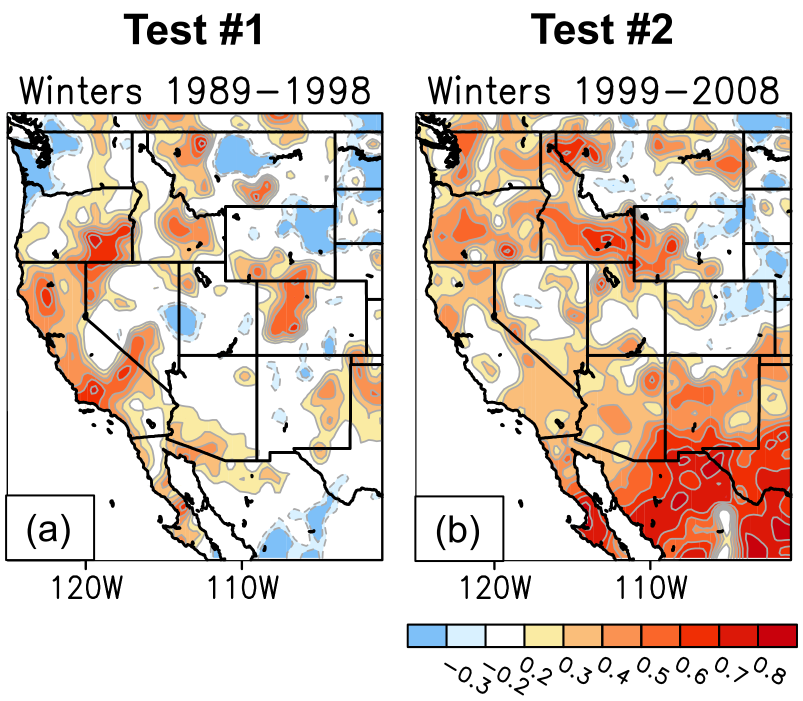
Part III: Hindcast skill of winter precipitation over the western U.S.

Part IV: Experimental seasonal forecast for winter 2021-22

Part V: Future plans: Statistical and Machine Learning applications

Precipitation hindcast skill

Hindcast skill during the past three decades



Positive correlation
→
Better hindcast skill

- Hindcast skill is assessed via *n-fold cross-validation* for a combination of predictor patterns.

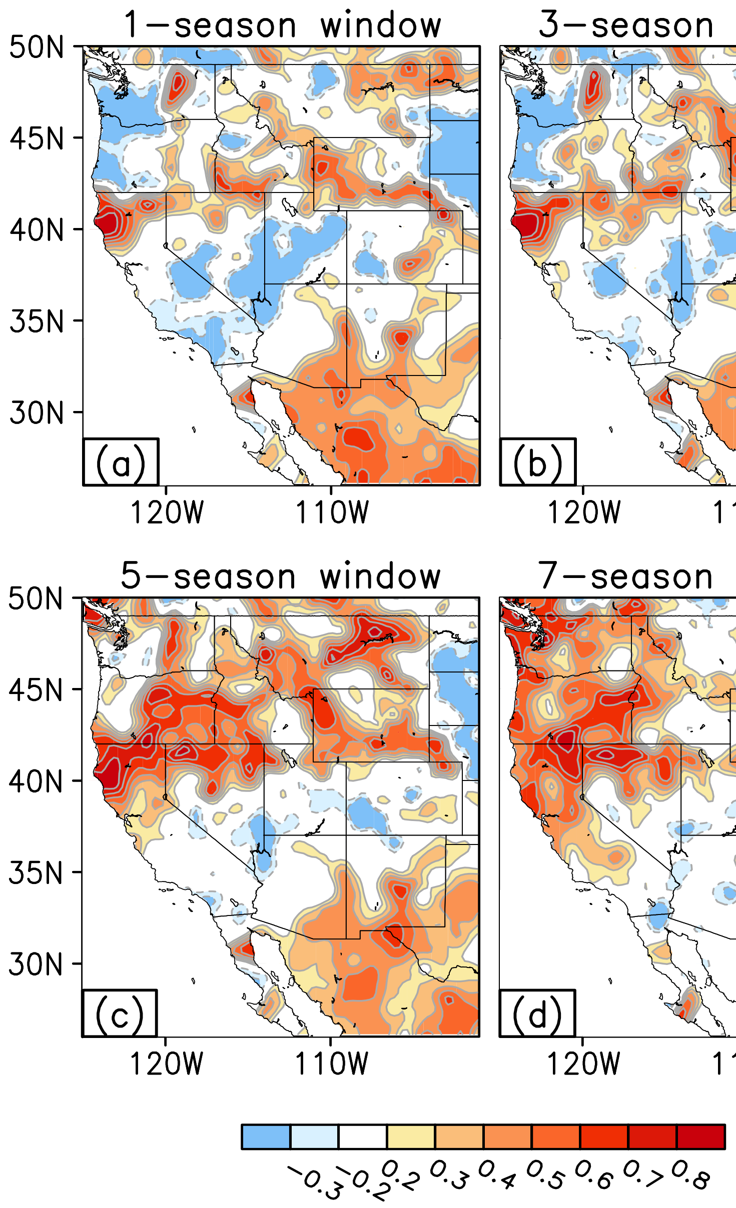
- The model is fit iteratively n times, each time training the data on *n-1* folds and evaluating on the the validation set.

Test #1	1949-58	1959-68	1969-78	1979-88	1989-98	1999-08	2009-18
Test #2	1949-58	1959-68	1969-78	1979-88	1989-98	1999-08	2009-18
Test #3	1949-58	1959-68	1969-78	1979-88	1989-98	1999-08	2009-18

Testing set
Training set

- Correlation coefficients between the hindcast and observed precipitation anomalies are displayed over individual test sets.

Skill score with change in length of predictor window



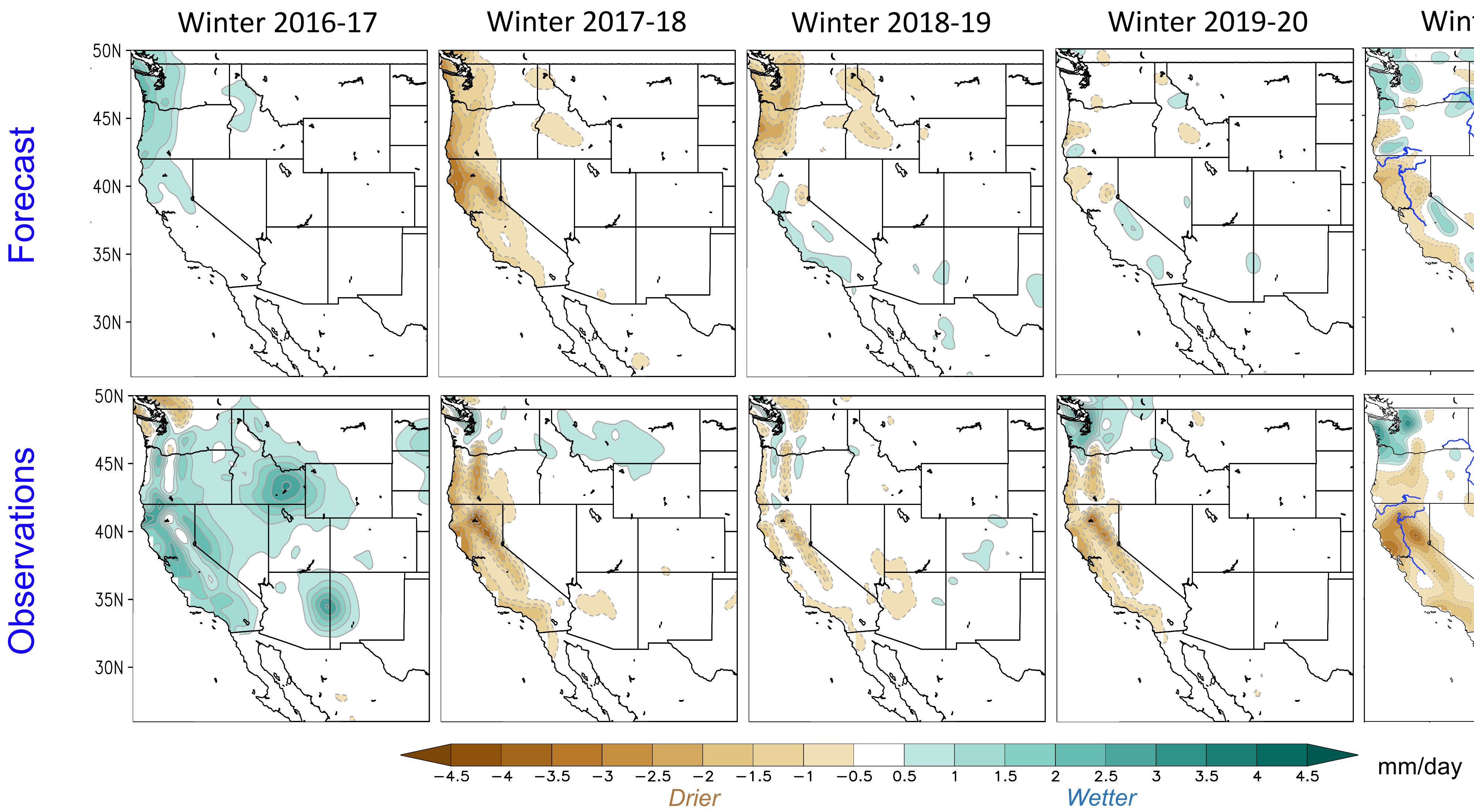
Precipitation hindcast skill

- ❑ Hindcast skill is assessed as a function of the length of temporal sampling window employed in an extended-EOF analysis.

Spring (t-6)	Summer (t-5)	Fall (t-4)	Winter (t-3)	Spring (t-2)	Summer (t-1)	Fall (t)	Winter (t+1)

- ❑ Skill assessment based on correlation coefficients values vis-à-vis observations
- ❑ Training period: 1948-2008 winters
- ❑ Validation period: 2009-2018 winters
- ❑ Using a longer temporal sampling window of predictors leads to better forecast skill

Past winter precipitation forecasts and verification



- Observed vs. predicted winter (December-February) precipitation anomalies over the Western U.S.
- Forecasts are generated from the modal contributions of Pacific SST PCs using a 5-season predictor window

Outline

Part I: Motivation and hypothesis

Part II: Evolution-centric statistical forecast technique

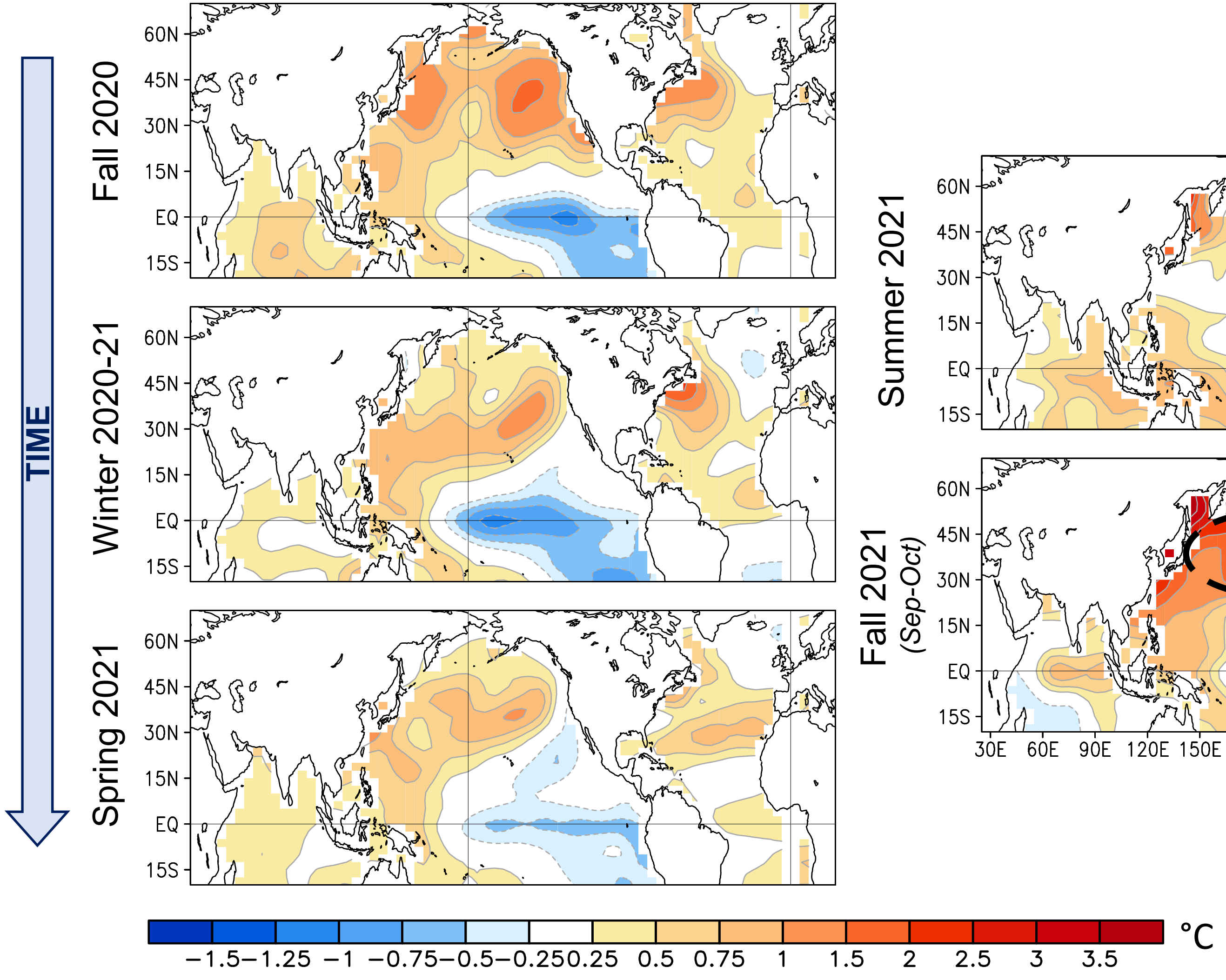
Part III: Hindcast skill of winter precipitation over the western U.S.

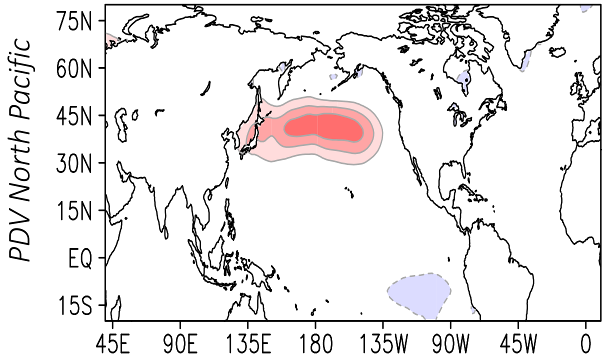
Part IV: Experimental seasonal forecast for winter 2021-22

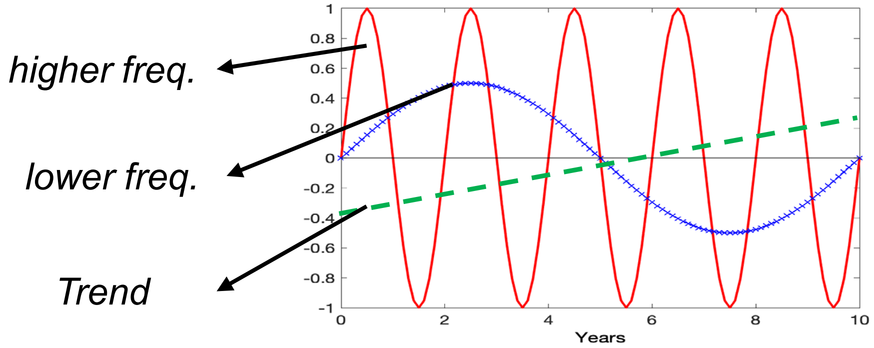
Part V: Future plans: Statistical and Machine Learning applications

Sea surface temperatures during the prior seasons

- Equatorial SSTs were below normal across most of the central and eastern Pacific Ocean.
- Persistent warm SSTs in the northern Pacific Ocean across multiple seasons.
- Interesting resemblance to the Pacific Decadal Variability: North Pacific mode (low frequency)



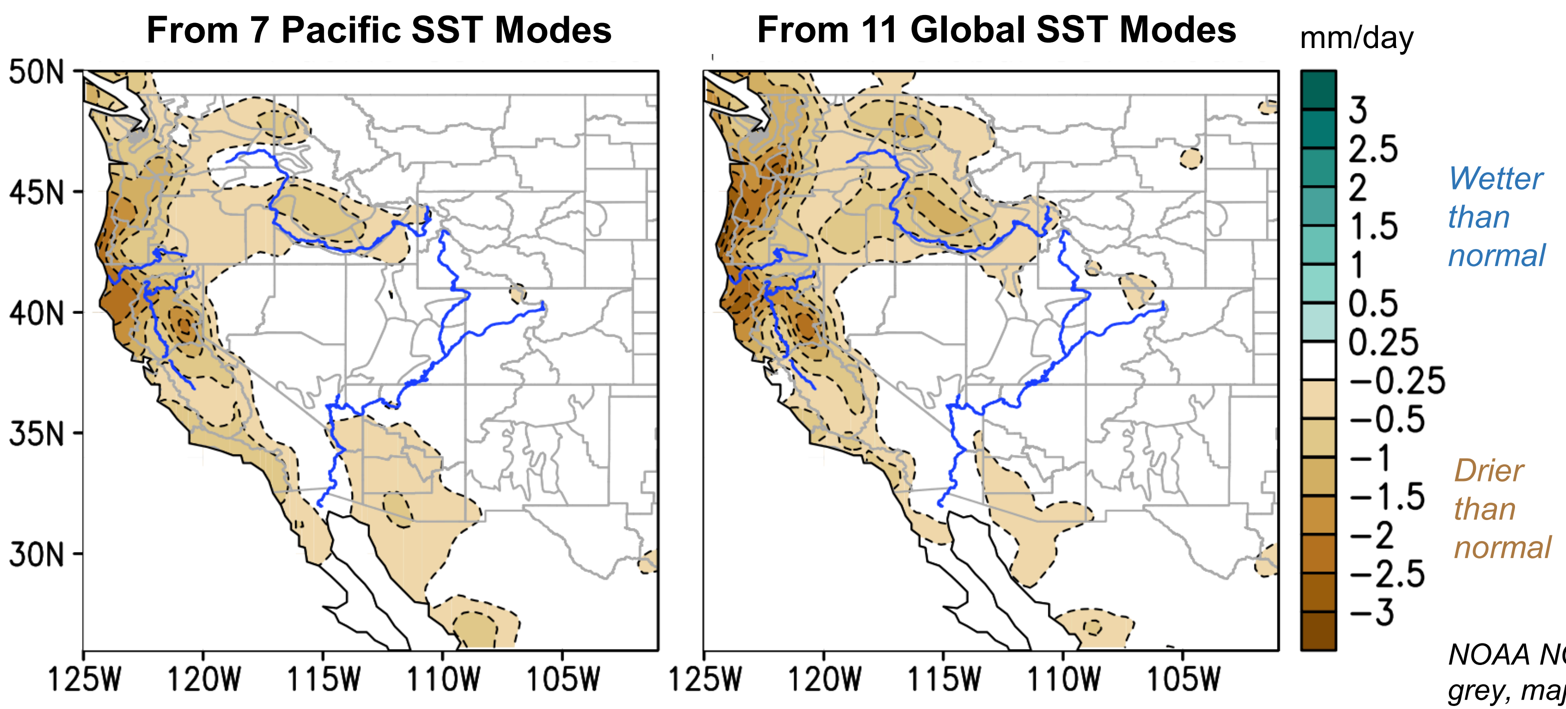




Experimental seasonal winter precipitation forecast *(Dec-Feb)*

2021-22 ***Dec-Feb*** **Forecast** *(issued 7 Nov '21)*:

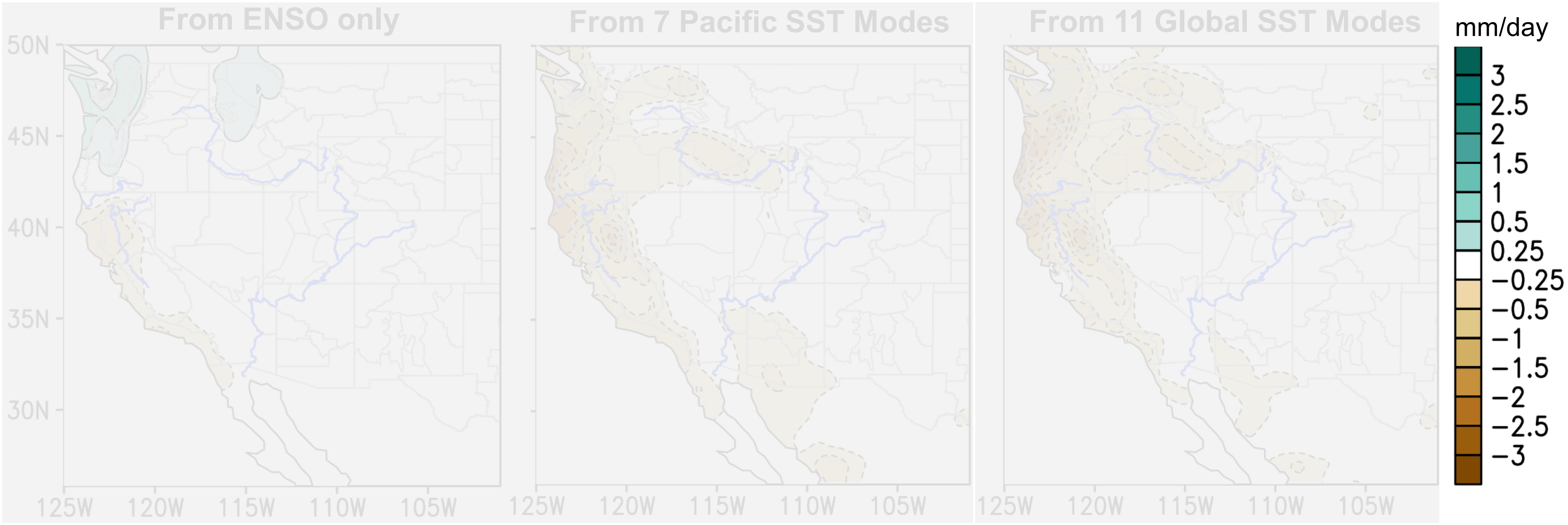
Seasonal Precipitation Anomalies



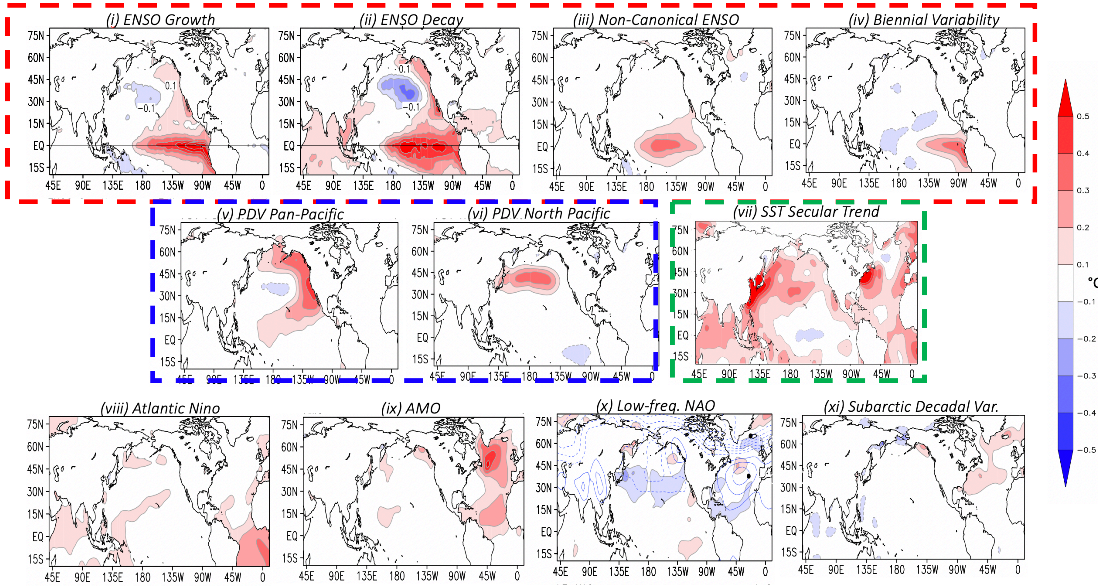
NOAA NCEI Climate Divisions outlined in grey, major western rivers in blue; Base period for precipitation anomalies: 1981-2010

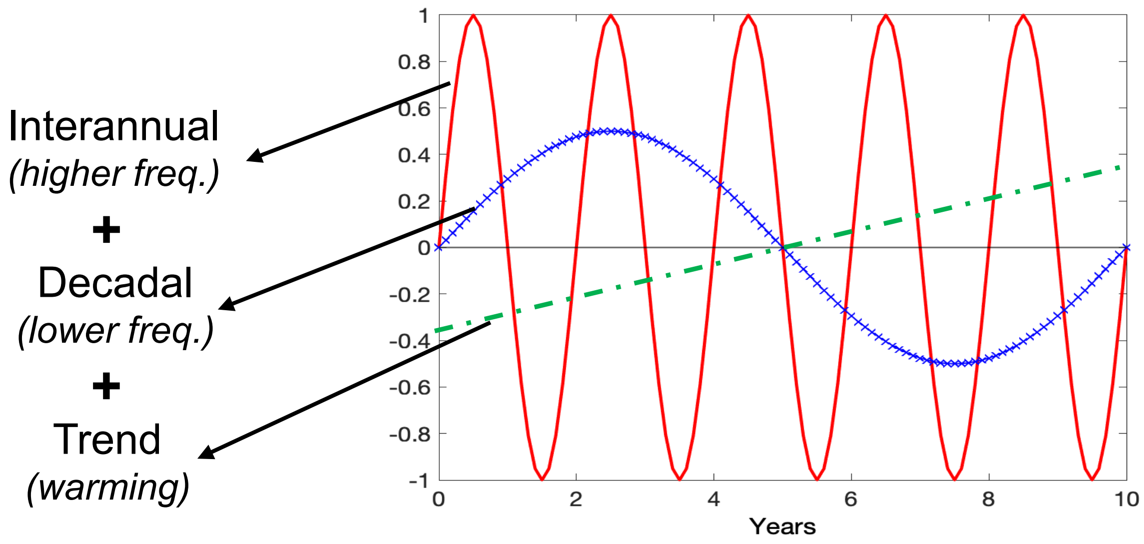
- Our experimental forecast favors *drier-than-normal* conditions in northern and southern California.
- *Near-normal* rainfall forecasted in the Upper Colorado river basin.

Modal attribution to the seasonal forecast



- Based on contribution from ENSO only, we have the classical dipole in precipitation footprint.
- However, our analysis emphasizes the need to consider modes of variability across a wide range of timescales.





Outline

Part I: Motivation and hypothesis

Part II: Evolution-centric statistical forecast technique

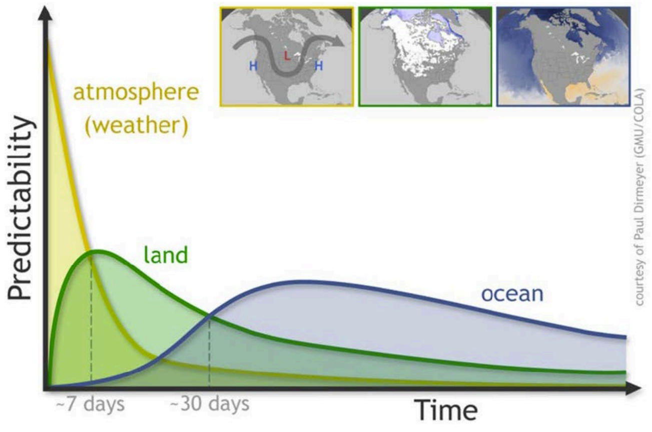
Part III: Hindcast skill of winter precipitation over the western U.S.

Part IV: Experimental seasonal forecast for winter 2021-22

Part V: Future plans: Statistical and Machine Learning applications

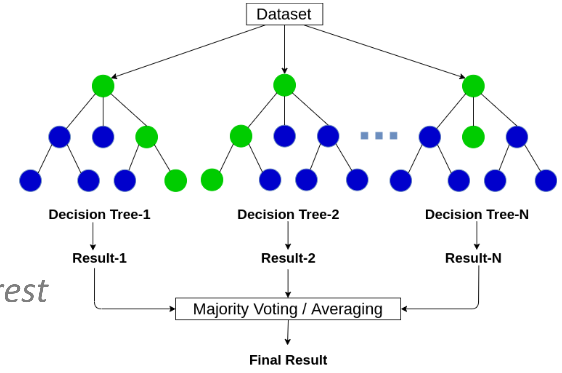
Future Work

- **Proposed application #1**: Expansion of current analysis to incorporate other sources of predictability, followed by exploration of their order of importance [submitted to *NASA ROSES A.34 Earth Science Applications: Water Resources* in Sep 2021]



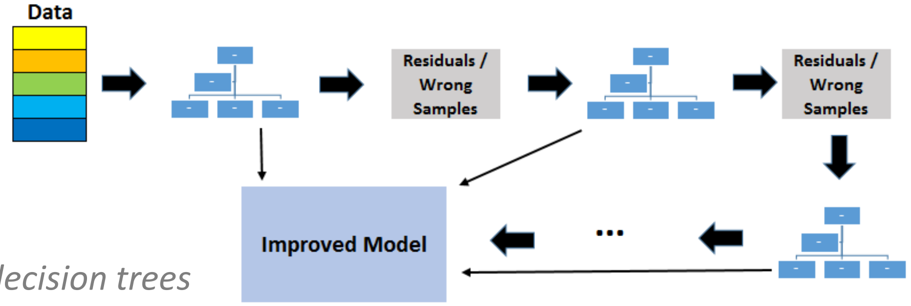
- **Advantages**:
  - Ability to accommodate *multiple sources of predictability* in a unified framework
  - Extracts clearly *delineated predictors* which are mutually orthogonal by construction
  - Obtains combination of *independent* predictors that maximizes the variance explained in a *dependent* predictand

- **Proposed application #2**: Identification of predictors (features) from observational analysis, then, training Machine Learning algorithms on the extracted patterns [*project currently underway with JPL ML group*]



Random Forest

- **Advantages**:
  - Training on actual observed patterns eliminates non-physicality
  - Can accommodate *non-linearity* in process interactions
  - Able to handle very large datasets and multiple variables



Gradient-boosted decision trees

Concluding remarks

- The observational record (including paleoclimate proxies) over the Western U.S. is characterized by prominent *low-frequency variability*.

- We demonstrate the need to accommodate the sources of predictability ranging from interannual to decadal-multidecadal timescales in context of longer lead seasonal forecasting.

- Regional hydroclimate predictions at longer lead times benefit from characterization of the *evolution* of the nascent and mature phases of variability.

- Based on the retrospective forecasts, thus far, global and basin-scale modes of SST variability are shown to be viable predictors of wintertime precipitation over the Western U.S.

- Our experimental forecast for winter (Dec-Feb) 2021-22 precipitation favors drier-than-normal conditions in northern and southern California, and near-normal conditions in the Upper Colorado river basin.

Thank you for listening!

Contact: agniv.sengupta@jpl.nasa.gov